**Store Sales Analysis Report**

**DEPI Graduation Project (May 2025)**

# **Executive Summary**

This report explores IKEA store sales data, covering orders, customers, products, and regions. It highlights key insights into sales trends, product performance, customer segments, and geographic patterns. These findings support data-driven strategies to boost sales, enhance customer targeting, and drive sustainable business growth.

# **2.Introduction & Goal**

**Introduction**

* This project presents a comprehensive analysis of IKEA store sales data with the objective of uncovering valuable business insights. The analysis focuses on key performance metrics to better understand customer behavior, assess product performance, and identify order patterns.
* The study investigates customer demographics, sales transactions, product categories, and order-related details to reveal the most influential factors contributing to overall sales performance.
* By exploring these dimensions, we aim to build a clear picture of how customers interact with the store’s offerings and what drives revenue growth across different segments and regions.

**The Goal**

The primary goal of this project is to generate actionable insights that can inform strategic decisions and support data-driven improvements.  
The analysis is designed to:

* Optimize sales strategies by identifying successful trends and performance gaps.
* Improve customer satisfaction through better understanding of customer behavior and preferences.
* Enhance overall business performance by supporting leadership with clear, evidence-based recommendations rooted in data.

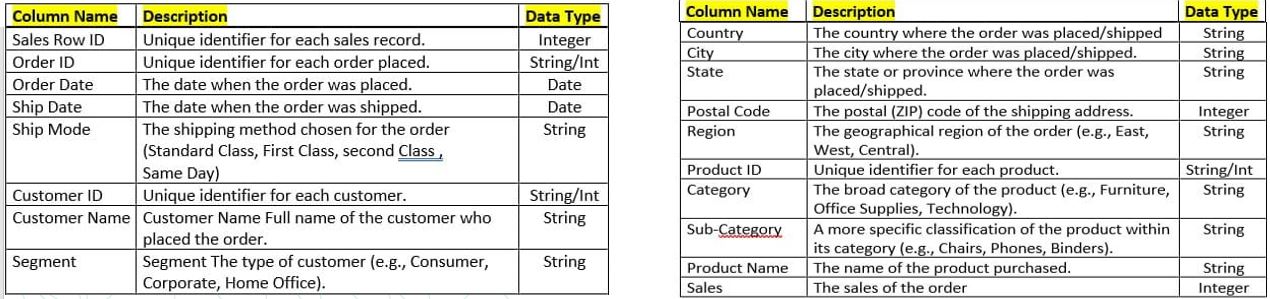
### **3.Steps Followed in the Analysis**

* Dataset Overview
* Cleaning and Preprocessing
* Data Modeling
* Questions Analysis
* KPIs and Business Insights
* Conclusion
* Recommendations

### **4- Dataset Overview**

The dataset contains 9,800 records and 18 columns related to IKEA store sales, including customer details, orders, products, and shipping. It combines both quantitative and qualitative data types. The goal of analyzing this dataset is to extract insights into customer behavior and sales performance to support business decisions.

### **5-Data Fields Description**



### **6- Data Cleaning & Preprocessing**

**Initial Cleaning**

* The dataset was initially found in a clean format, requiring minimal cleaning efforts.
* No duplicated rows were detected, indicating good data consistency.
* The "Postal Code" column had 11 missing values, which were filled using appropriate postal codes based on the corresponding state.

**Transforming Date Fields**

* The "Order Date" column was split into three separate columns: "order-Day" , "order-Month" and "order-Year".
* Similarly, the "Ship Date" column was divided into: "ship-Day", "ship-Month", and "ship-Year".
* A new column named "Ship-Quarter" was created based on the ship month to facilitate quarterly analysis.
* Both "Order Date" and "Ship Date" formats were standardized for consistency.
* A "Duration" column was added by calculating the difference between the "Ship Date" and the "Order Date".

**Column Handling**

* All existing columns were preserved as they were relevant to the analysis.
* No columns were dropped during the preprocessing phase.

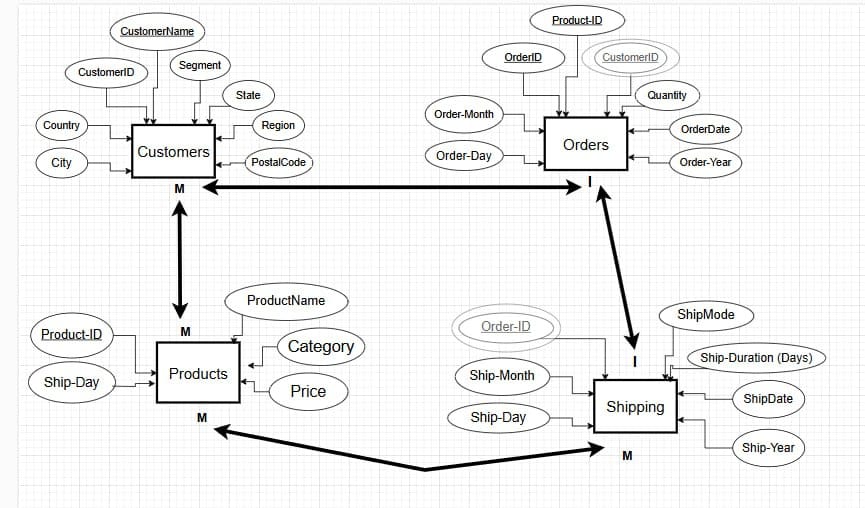
### **7- Data Modeling**

To ensure a structured and efficient analysis, we began by exploring the dataset thoroughly to understand its content and relationships. The data was logically divided into four main sheets:

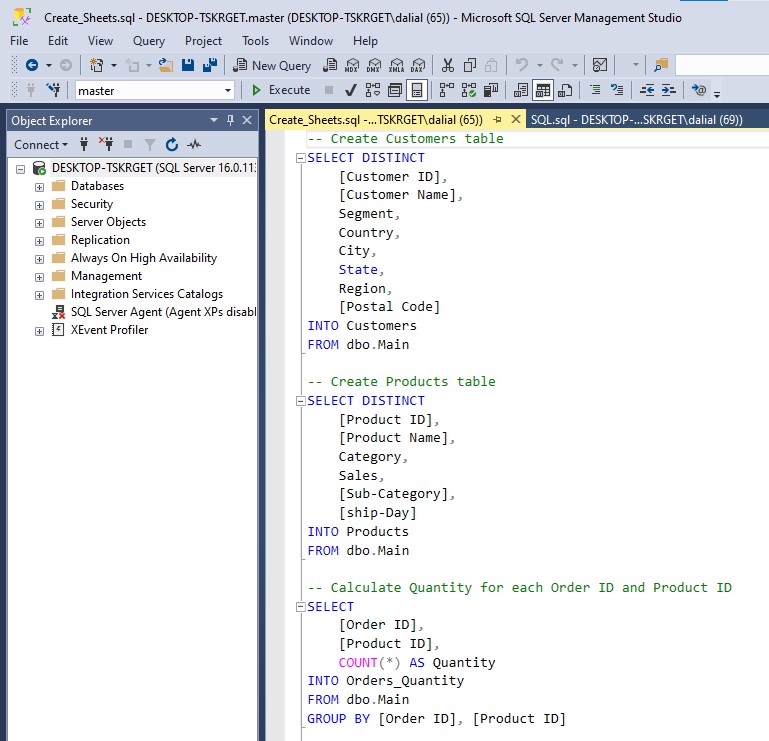
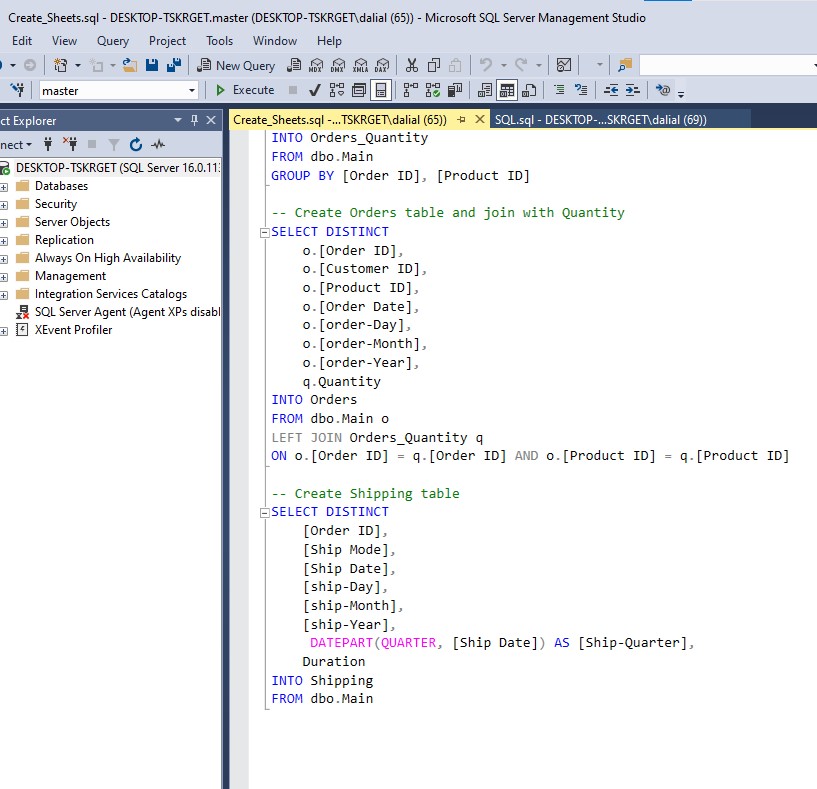
* **Customers Sheet:** Included all relevant customer information.
* **Orders Sheet:** Contained order details and transaction records.
* **Products Sheet:** Focused on product descriptions and classifications.
* **Shipping Sheet:** Captured shipping details and logistics data.

We then established clear relationships between these sheets to maintain data integrity and enable cross-analysis. To visualize the data structure and its interconnections, we created an Entity-Relationship Diagram (ERD) using "*Figma"*. All data modeling and transformation steps were executed using **SQL**, ensuring accuracy, consistency, and scalability of the process.

**ERD Diagram**



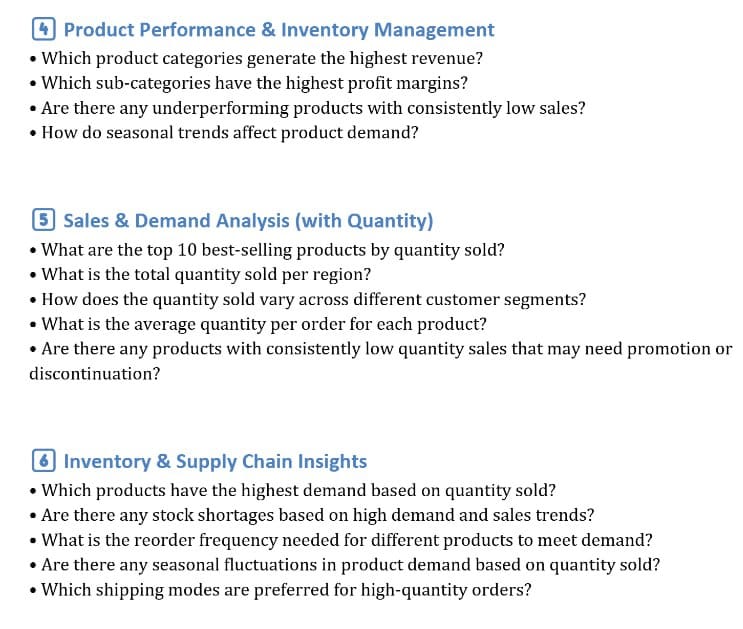
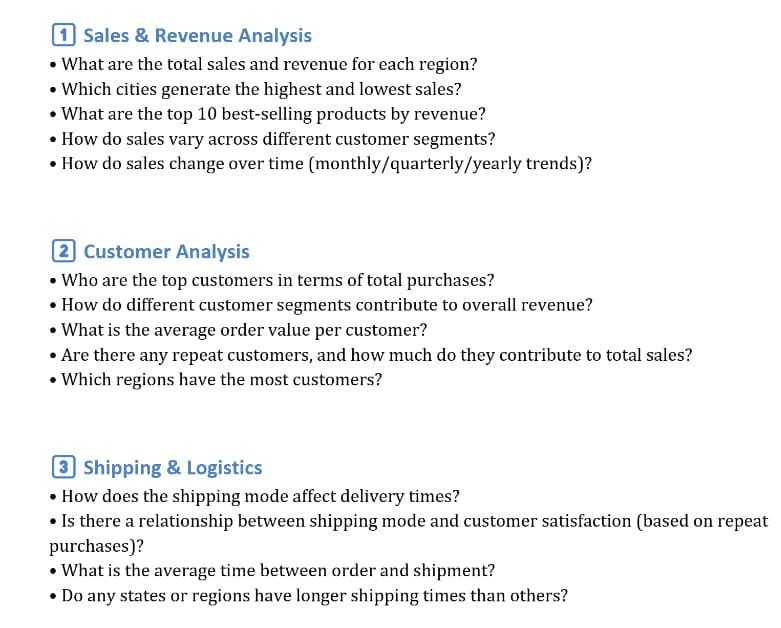
**Modeling Using SQL**

### **8. Question Analysis**

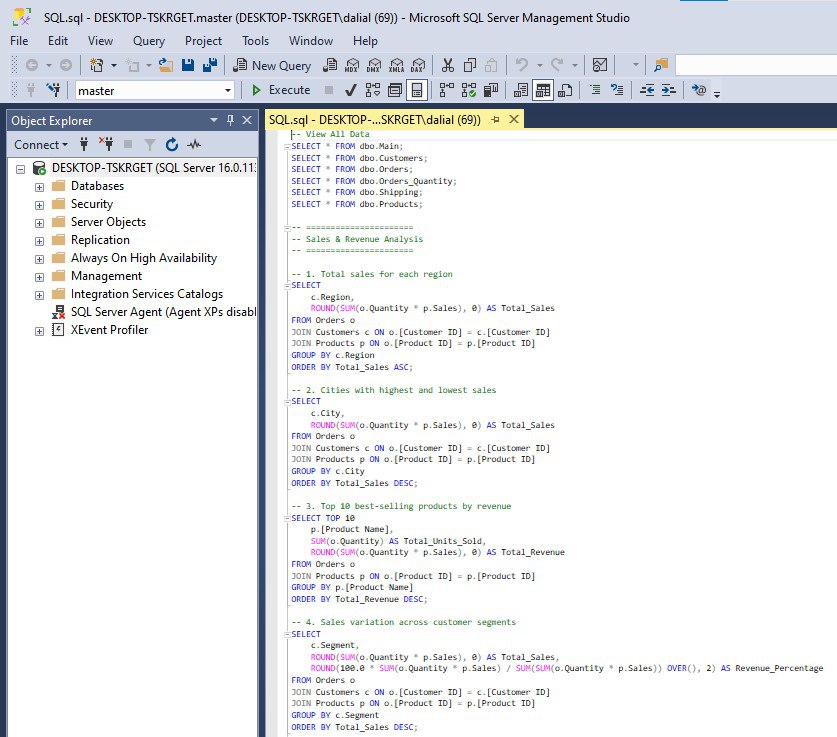
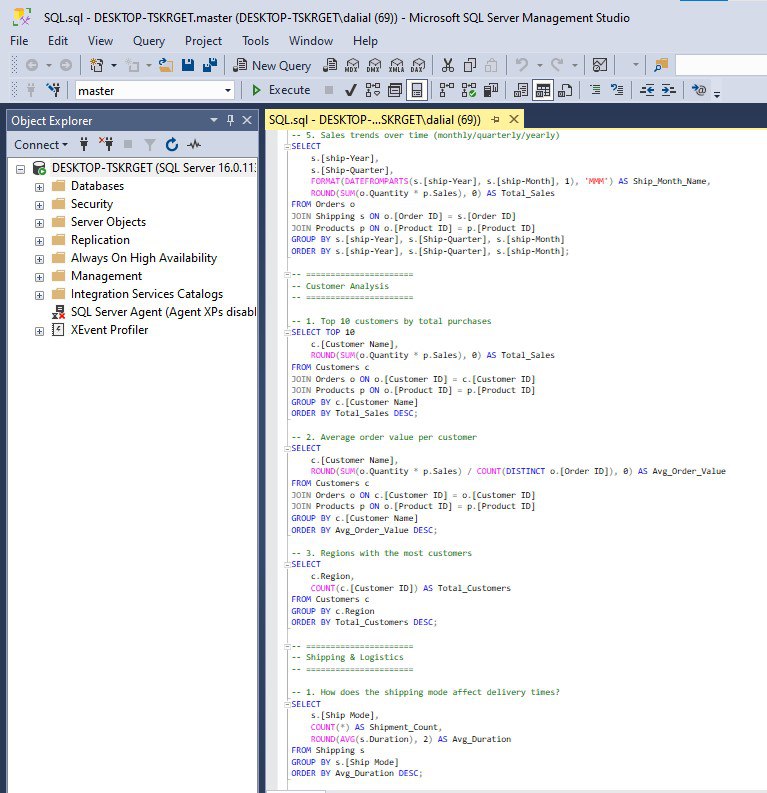
* We identified and defined key business questions relevant to customer behavior, product performance, and order patterns.
* The questions were carefully extracted based on the available data to ensure they could be answered effectively.
* We used SQL queries to analyze and extract the necessary insights from the dataset.
* Then, we designed interactive dashboards to visually represent the answers, making them clear and actionable for decision-makers.

**Business Questions**



**Answering Using SQL**

This is a part of answering questions:



### **9- Insights and KPIs Derived from Dashboards**

**Customer Behavior Dashboard**

General KPIs:

* **Total Customers: 793**
* **Percent of Repeat Customers: 98% (Very High)**
* **Average Orders per Customer: 6**
* **Average Sales per Customer: 21K**

General insights:

* The **Consumer segment** not only has the highest number of customers but also contributes the most to total sales, reinforcing its strategic importance.
* A steady **growth in customer base over the years** suggests increasing brand reach and market penetration.
* The **Consumer-dominant** region shows the highest customer density , while other regions have significantly lower customer count.

**Orders Dashboard**

General KPIs:

* **Total Orders:** 10K
* **Average Orders per Month:** 465
* **Average Orders per Day:** 8
* **Total Sales:** 16 million

General insights:

* The **Consumer** segment is the largest contributor to order volume, making it the primary target for marketing and product strategies.
* **Office Supplies** dominate sales volume, suggesting high demand and possibly repeat purchases. Technology, while valuable, has lower order frequency.
* There are noticeable seasonal trends, with a spike in Q3. Promotional campaigns could be optimized during low-performing months.
* **California** is the clear market leader. Other states like **Texas** and **New York** also represent strong markets, indicating regional expansion opportunities.

**Products Dashboard**

General KPIs:

* **Total Products: 1.9K**
* **Total Quantity: 11.3K**
* **Number of Sub-Categories: 17**
* **Avg. Quantity Per Order: 2**

General insights:

* **Office Supplies** have the highest number of orders and the widest range of products, indicating consistent demand and customer preference for this category.
* Despite having fewer orders, the **Chairs** sub-category generates the highest revenue, highlighting it as a key high-value product.
* The quantity of products ordered peaks in **October and November**, suggesting seasonal spikes likely due to promotions or end-of-year demand.
* An average of 2 items per order indicates frequent multi-item purchases, offering opportunities for bundling or upselling strategies.

### **10- Conclusion**

* **Consumer Segment Dominates**: The Consumer segment leads in both customer count and total sales, making it the most critical area for growth.
* **Product Performance**: Office Supplies generate the highest order volume, while Chairs, within Furniture, contribute the most revenue, indicating their high value.
* **Seasonal Trends**: There are clear peaks in order quantities around October and November, suggesting a seasonal demand pattern.
* **Customer Growth**: The customer base has shown a steady increase from 2015 to 2018, indicating successful brand reach and market penetration.
* **Geographical Distribution**: California, Texas, and New York are key markets, with California having the highest concentration of customers.
* **Multiple Item Purchases**: Customers tend to buy more than one product per order, opening up opportunities for cross-selling and bundling strategies.

### **10- Recommendations**

#### **Sales Strategy**

* **Focus on Consumer Segment**: Since the Consumer segment contributes the most to sales and customer volume, tailor promotions, loyalty programs, and product offerings specifically for this segment.
* **Boost Underperforming Segments**: Develop targeted marketing campaigns to improve engagement from the Home Office and Corporate segments.
* **Capitalize on Best-Selling Sub-Categories**: Promote high-performing sub-categories like Chairs, which generate strong revenue, and explore bundling them with related items.

#### **Product Strategy**

* **Expand Office Supplies Variety**: Given their high order volume, consider diversifying the Office Supplies category to maintain dominance.
* **Monitor Technology Sales**: Technology has lower order frequency—evaluate product offerings, pricing, or visibility to boost its performance.
* **Use Quantity Trends**: High purchase quantities during months 10 and 11 suggest promotional campaigns should be intensified during this period.

#### **Customer Engagement**

* **Leverage Repeat Customers**: With a 98% repeat customer rate, implement loyalty and referral programs to further boost retention.
* **Engage Top Customers**: Recognize and reward top spenders like Joel Eaton and Zuschuss Carroll through VIP programs or exclusive offers.
* **Improve Low-Performing Regions**: Regions with low customer counts could benefit from localized marketing or partnerships to improve presence.

#### **Geographical Expansion**

* **Strengthen in High-Performing States**: States like California, New York, and Texas show strong order activity—invest more in regional promotions and logistics there.
* **Explore Untapped Regions**: Use customer distribution data to identify underperforming areas and assess feasibility for market entry or campaign testing.

#### **Time-Based Strategy**

* **Plan for Seasonal Peaks**: Increase stock and staffing ahead of peak order months (October, November).
* **Boost Low-Season Sales**: Introduce discounts or new product launches during slower months (January, February).

#### **Operational Improvements**

* **Optimize Shipping Duration**: Analyze the newly created “Duration” field to reduce delivery times and improve customer satisfaction.
* **Ensure Data Quality**: Continue data validation processes, especially for critical fields like postal codes and date entries.
* **Enhance Data Modeling**: Maintain clear ERDs and logical data separation (as done in SQL modeling), ensuring scalable dashboards and future updates.